

# buscar Raíces con Matrices Colega?

Andrés Hernández, Francisco Díaz  
y Joaquín Peñuela

## 1. Introducción

Históricamente, el problema de calcular raíces de polinomios siempre ha sido de gran interés debido a sus amplias aplicaciones. En primera instancia, algunos podrían pensar que no hay necesidad de querer conocer el cero de un polinomio, o de una función cualquiera; sin embargo, conocer estos ceros tiene muchas aplicaciones, desde poder optimizar sistemas en economía hasta analizar cuando un material podría sufrir una transición de fase.

Es bien conocido el hecho de que para polinomios de grado superior a 4 no existe una fórmula analítica para calcular sus ceros. Por esta razón, se recurre a métodos numéricos. En general, existen diversos métodos para calcular las raíces de manera aproximada. Estos permiten calcularlos con una precisión tan pequeña como se desee a cambio de aumentar el número de iteraciones, lo cual implica un mayor tiempo de cómputo. Adicionalmente, la mayoría de estos métodos tienen en común que se desarrollan en la *base de monomios* del espacio polinomial. Es decir, que todo polinomio se piensa como una combinación lineal del conjunto  $\{x^k\}_{k=0}$ . La intuición nos muestra que en primera instancia no habría problema alguno en calcular las raíces usando esta base. Sin embargo, Wilkinson descubrió en 1950 que trabajar con esta base de polinomios es un problema *mal condicionado*.

En un principio, Wilkinson se preguntó qué tan grande sería el cambio en las raíces de un polinomio al ser modi-

ficado con una perturbación muy pequeña. La respuesta que encontró es que las raíces del polinomio cambiarían en órdenes de magnitud mucho mayores al orden de la perturbación, lo cual indicaba que se trataba de un problema mal condicionado. Para entender esto, defínase el polinomio de Wilkinson como  $P_{\text{Wilkinson}}(x) = \prod_{i=1}^{20} (x - i)$  y considérese el polinomio resultante dada una pequeña perturbación  $\varepsilon > 0$  dado por  $P(x) = P_{\text{Wilkinson}}(x) + \varepsilon x^{14}$ .

En la figura 1, al tomar distintos valores de la perturbación en órdenes de  $10^{-10}$ , se evidencia que la perturbación en las raíces es de órdenes mucho mayores. Más aún, estas raíces dejan de ser en su totalidad números reales y pasan a tener una componente imaginaria.

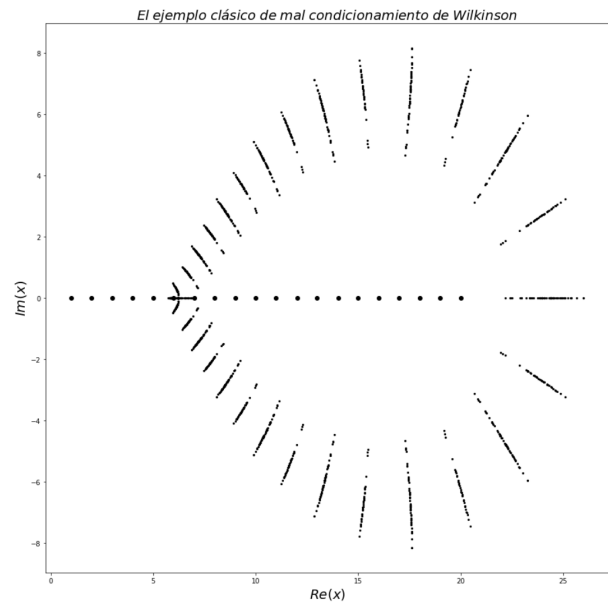


Figure 1: Raíces del polinomio de Wilkinson bajo perturbaciones.

Universidad de los Andes. Departamento de matemáticas.

[a.hernandez12@uniandes.edu.co](mailto:a.hernandez12@uniandes.edu.co)

[f.j.diaz@uniandes.edu.co](mailto:f.j.diaz@uniandes.edu.co)

[j.penuela@uniandes.edu.co](mailto:j.penuela@uniandes.edu.co)

El código de la implementación utilizada se puede encontrar en el siguiente repositorio:

[https://github.com/PachoDp13/Personal/tree/main/Proyecto\\_Analisis\\_Numerico](https://github.com/PachoDp13/Personal/tree/main/Proyecto_Analisis_Numerico)

Esto indica que perturbaciones muy pequeñas pueden producir cambios drásticos en las raíces. De esta forma se puede ver que el problema de calcular las raíces de un polinomio en *función de sus coeficientes* está mal condicionado cuando se considera la base de monomios. No obstante, existe otra base de polinomios en donde el problema de buscar raíces *en función de los valores del polinomio* se convierte en un problema de valores propios, el cual está bien condicionado. Esta base se conoce como la base de *polinomios de Chebyshev*.

La idea general de traducir el problema de encontrar raíces al cálculo de valores propios es construir una matriz que tenga como polinomio característico el polinomio del cual queremos conocer las raíces. Al trabajar con la base monomial, dado un polinomio  $P(x) = c_0 + c_1x + \dots + c_{n-1}x^{n-1} + x^n$  estas matrices son conocidas como *matrices compañeras* y lucen de la siguiente manera:

$$C_p = \begin{bmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{n-1} \end{bmatrix} \quad (1)$$

Sin embargo, como se mencionó anteriormente, para evitar el mal condicionamiento característico de la base monomial, es necesario utilizar otro tipo de matrices que se construyen a partir de la base de Chebyshev, de ahí el título de este documento.

## 2. Polinomios de Chebyshev

Los polinomios de Chebyshev son un conjunto de polinomios ortogonales que aparecen de manera natural como solución de la ecuación diferencial de Chebyshev dada por:

$$(1 - x^2) \frac{d^2 y}{dx^2} - x \frac{dy}{dx} + p^2 y = 0 \quad (2)$$

Donde  $p$  es una constante de valor real o complejo. Las soluciones de esta ecuación diferencial se pueden encontrar por medio de series de potencias de la forma  $y = \sum_{n=0}^{\infty} a_n x^n$ , donde los coeficientes cumplen la relación de recurrencia dada por:

$$a_{n+2} = \frac{(n-p)(n+p)}{(n+1)(n+2)} a_n \quad (3)$$

Cabe notar, que esta serie converge en el intervalo dado por  $|x| < 1$ , por esto se tiene que los polinomios son ortogonales en ese intervalo. Esta propiedad de ortogonalidad se cumple para el producto interno definido de la siguiente manera:

$$\langle T_n(x), T_m(x) \rangle = \int_{-1}^1 \frac{T_n(x) T_m(x)}{\sqrt{1-x^2}} dx \quad (4)$$

Para el cual se tiene que:

$$\langle T_n(x), T_m(x) \rangle = \begin{cases} 0 & \text{si } n \neq m \\ \pi & \text{si } n = m = 0 \\ \pi/2 & \text{si } n = m \neq 0 \end{cases} \quad (5)$$

Cabe resaltar que estos polinomios forman una base, por lo que cualquier polinomio se puede escribir como una combinación lineal de ellos  $p(x) = \sum_k a_k T_k(x)$ . Esto se conoce como la representación de Chebyshev, donde los coeficientes están dados simplemente por la proyección del polinomio respecto a la base:

$$a_k = \langle f(x), T_k(x) \rangle$$

Así, si se desea proyectar un polinomio  $p(x)$  en esta base se debe utilizar una matriz de cambio de base que permita convertir los coeficientes de la base monomial a los coeficientes de la base de Chebyshev.

**Matriz de cambio de base:** Dado un polinomio  $p_n(x)$  definido sobre  $x \in [a, b]$ , realizamos un reescalamiento de la forma  $y \equiv \frac{2x-(b+a)}{b-a}$  tal que:

$$p_n(x) = \sum_{k=0}^n a_k T_k(y) = \sum_{k=0}^n b_k y^k, \quad y(x) \in [-1, 1] \quad (6)$$

En la implementación del cambio de base, es posible disminuir a la mitad el costo de transformar de  $\{a_j\}$  a  $\{b_j\}$ . Para esto, estudiamos por separado los coeficientes pares e impares:

$$b_{even} = Q^{even} a_{even} \quad b_{odd} = Q^{odd} a_{odd} \quad (7)$$

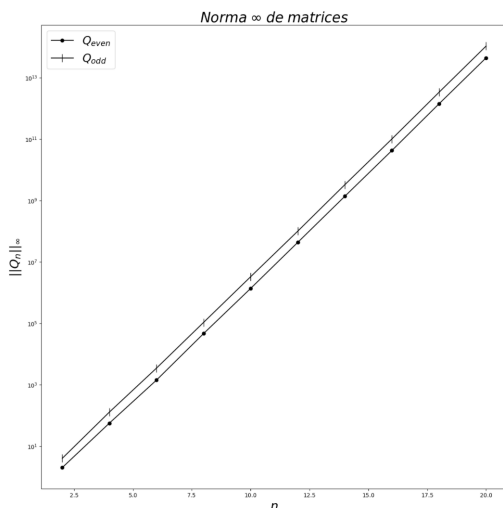
Donde los elementos de estas matrices de cambio de base son obtenidos recursivamente al utilizar la estructura de los polinomios de Chebyshev:

$$Q_{11}^{even} = 1, \quad Q_{jj}^{even} = 2^{2j-3}, \quad Q_{11}^{odd} = 2^{2j-3}$$

$$Q_{j-k,j}^{even} = \left[ -\frac{(2j-2k)(2j-2k-1)}{2k(4j-2k-4)} Q_{j-k+1,j}^{even} \right]$$

$$Q_{j-k,j}^{odd} = \left[ -\frac{(2j-2k+1)(j-k)}{k(4j-2k-2)} Q_{j-k+1,j}^{odd} \right]$$

Un problema que surge al trabajar con estas matrices de cambio de base, y que explica de dónde proviene el mal condicionamiento de la base monomial, es que la norma de estas crece exponencialmente en función del grado del polinomio.



**Figure 2:** Norma  $\|\cdot\|_\infty$  de las matrices  $Q^{even}$  y  $Q^{odd}$  en función del grado del polinomio.

Particularmente, en la figura 2 se muestra cómo la norma  $\infty$  de las matrices de cambio de base crece exponencialmente, aproximadamente como [2]:

$$\|Q^{even}\|_\infty \sim 0.016(5.8)^n \quad \|Q^{odd}\|_\infty \sim 0.39(5.8)^n \quad (8)$$

El siguiente problema que surge es, ¿cómo calcular la expansión de Chebyshev para una función arbitraria  $f$ ? Para solucionar esto, al recordar la relación de ortogonalidad 5, es natural pensar en realizar una interpolación. En esta,  $f(x)$  debe evaluarse en un conjunto de puntos discretos en el intervalo objetivo (nodos de Chebyshev):

$$x_j = \frac{b-a}{2} \cos((j+1/2)\pi/(n+1)) + \frac{b+a}{2} \quad (9)$$

Luego, los coeficientes de la expansión 6 estarán dados por la multiplicación del vector que contiene el valor de  $f$  evaluada en estos puntos y una matriz cuyas entradas son simples funciones trigonométricas (polinomios de Chebyshev evaluados en los nodos de la rejilla):

$$a_k = \frac{2}{n+1} \sum_{j=0}^n f(x_j) \cos((j+1/2)\pi/(n+1)) \quad (10)$$

Así, es claro que cualquier polinomio  $P(x)$  o función  $f(x)$  se puede escribir respecto a la base de Chebyshev. Luego, el siguiente paso es buscar las raíces de la matriz asociada a esta base polinomial: *la matriz colega*. Para esto, enunciamos el siguiente teorema:

**Teorema [7]:** Sea  $P(x)$  un polinomio expresado en la base de Chebyshev:

$$P(x) = \sum_{k=0}^n a_k T_k(x), \quad a_n \neq 0 \quad (11)$$

Entonces sus raíces (con multiplicidades incluidas) son los valores propios de la matriz colega definida de la siguiente forma:

$$C = \begin{pmatrix} 0 & 1 & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ & \frac{1}{2} & 0 & \frac{1}{2} & \\ & & \ddots & \ddots & \ddots \\ & & & \frac{1}{2} & 0 \end{pmatrix} - \frac{1}{2a_n} \begin{pmatrix} & & & & \\ & & & & \\ & & & & \\ & & & & \\ a_0 & a_1 & a_2 & \dots & a_{n-1} \end{pmatrix}$$

**Figure 3:** Matriz Colega del polinomio  $P(x)$ .

Así, para calcular los ceros de cualquier polinomio simplemente hay que definir su matriz colega y calcular sus valores propios. Así, hemos reducido el problema de buscar raíces a un cálculo de valores propios. Esto se puede utilizar con la amplia maquinaria que tenemos del álgebra lineal. Particularmente, para nuestro caso esto lo solucionaremos utilizando el algoritmo conocido como *iteración QR*.

### 3. Iteración QR

Dada una matriz  $A \in \mathbb{R}^{n \times n}$  arbitraria, existe una matriz  $Q$  ortogonal y una matriz  $R$  triangular superior tales que  $A = QR$ . Estas matrices  $Q$  y  $R$  se pueden calcular a través de un método *bien condicionado* conocido como triangulación de Householder.

Utilizando esta descomposición, el algoritmo de la iteración QR para una matriz  $A$  es el siguiente:

1. Calcular la factorización QR de  $A$ , denotemos  $A_k = A$  y su factorización QR como  $A_k = Q_k R_k$ .
2. Definir  $A_{k+1} = R_k Q_k$ .
3. Repetir los pasos anteriores con  $A_{k+1}$  como punto de partida.

Para ver la relación que el algoritmo  $QR$  tiene con el cálculo de valores propios es necesario recordar la descomposición real de Schur:

**Teorema de la descomposición real de Schur :** Sí  $A \in \mathbb{R}^{n \times n}$  entonces existe una matriz ortogonal  $Q$  tal que

$$Q^T A Q = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1m} \\ & R_{22} & \cdots & R_{2m} \\ & & \ddots & \vdots \\ & & & R_{mm} \end{bmatrix} = T, \quad (12)$$

es una matriz cuasi-triangular superior semejante a  $A$  donde la diagonal está formada por bloques de

matrices  $2 \times 2$  correspondientes a las parejas conjugadas de valores propios complejos de  $A$  y bloques de matrices  $1 \times 1$  corresponden a los valores propios reales de  $A$ .

Una observación directa que se puede realizar, es que los valores propios de  $A$  y  $T$  son los mismos. Esto debido a que el determinante de una matriz cuasi-triangular superior es el producto del determinante de las matrices de la diagonal. Por lo que analizar los valores propios de  $T$  se reduce a analizar  $\det(T - \lambda I) = \prod_{i=1}^n \det(R_{ii} - \lambda_i I_{ii}) = 0$  donde  $I_{ii}$  es una matriz identidad con las mismas dimensiones que  $R_{ii}$ . Así, es claro que si  $R_{ii}$  es una matriz  $1 \times 1 \Rightarrow \lambda_i = R_{ii}$  (esto ocurre para los valores propios reales), mientras que si  $R_{ii}$  es de tamaño  $2 \times 2$  el valor propio  $\lambda_i$  será la solución al sistema  $\det(R_{ii} - \lambda_i I_{ii}) = 0$  (esto es lo que ocurre para los valores propios complejos). [1]

La relación entre la descomposición real de Schur y la iteración  $QR$  está en que la matriz a la que converge la iteración  $QR$  corresponde a la matriz de la descomposición real de Schur cuando el subíndice  $k$  es lo suficientemente grande. Debido a esto, calcular los valores propios de las matrices colega se reduce a aplicar la iteración  $QR$  y posteriormente analizar los bloques de la diagonal de la matriz resultante.

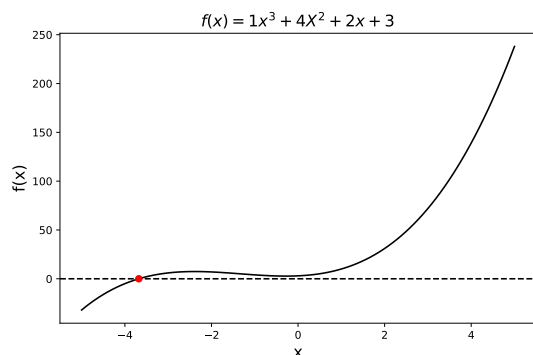


Figure 4: Búsqueda de raíces de Chebyshev.

A manera de ejemplo, en la anterior gráfica se ve que el cálculo de las raíces reales de un polinomio funciona de la manera adecuada. Aun así, se quiere entender qué ocurre con las raíces complejas. En la figura 4 solo graficamos la raíz real encontrada, no obstante, se encontraron dos raíces complejas, lo cual concuerda con los que resultan al utilizar métodos ya implementados en *numpy*. Más aún, al alterar un poco el algoritmo  $QR$ , es posible calcular estas raíces complejas a partir de los bloques  $2 \times 2$  de  $A_k$ , el cual es un problema analítico.

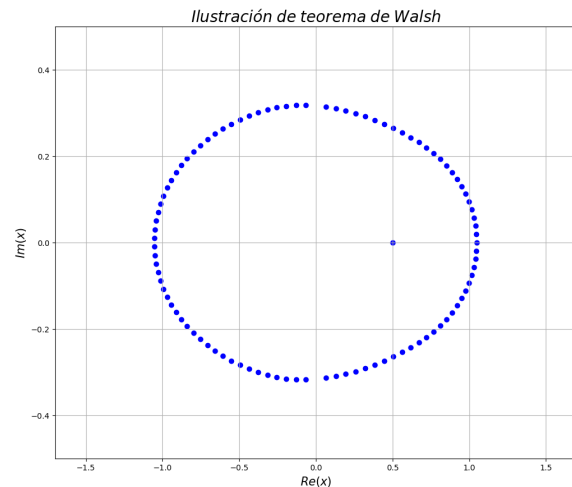


Figure 5: Raíces del polinomio interpolador de grado 100 de la función  $f(x) = \frac{x-1/2}{1+10x^2}$ .

Por otro lado, debido a que en las interpolaciones estamos utilizando polinomios de grado mucho mayor al de las funciones objetivos, las demás raíces (las que no estamos buscando) tienen un comportamiento interesante. Específicamente, al tomar la función  $f(x) = \frac{x-1/2}{1+10x^2}$  y realizar una interpolación de grado 100, en la figura 5 se evidenció que las raíces complejas se agrupan sobre una frontera, conocida como la *elipse de analiticidad de Bernstein*. Este fenómeno, en el cual las raíces de la aproximación se agrupan a lo largo de la máxima elipse de Bernstein es un caso especial de un teorema de Walsh [8].

Si bien este fenómeno es de interés, supera el alcance de este proyecto. Así, como en este punto es claro que el método propuesto al usar matrices colega funciona procedemos a estudiar el condicionamiento de este.

## 4. Condicionamiento

Dado un polinomio escalar mónico en la base de Chebyshev  $p(x)$ , es posible demostrar que si las raíces de  $p(x)$  se calculan como los valores propios de una matriz colega utilizando un algoritmo de valores propios estable regresivamente (como el algoritmo  $QR$ ), entonces las raíces calculadas son las raíces exactas de un polinomio mónico en la base de Chebyshev  $\tilde{p}(x)$  tal que:

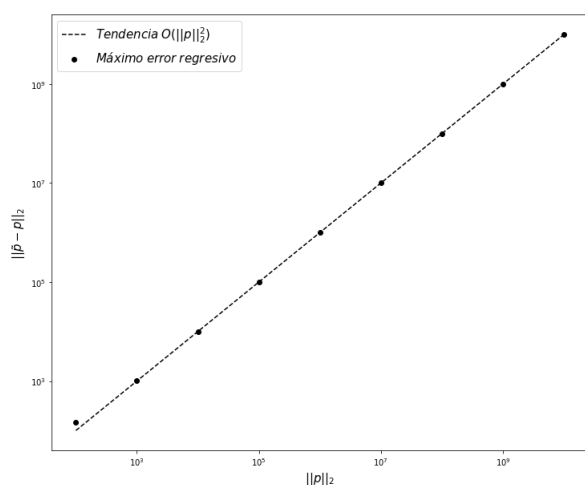
$$\frac{\|\tilde{p} - p\|_2}{\|p\|_2} = O(u)\|p\|_2 \quad (13)$$

La demostración de esto hace uso de una expansión a primer orden de la matriz colega de  $\tilde{p}$  con respecto a la perturbación del polinomio original  $p(x)$  y una descomposición en *matrices de Clenshaw* para posteriormente utilizar propiedades analíticas de las normas [5]. Más

aún, es posible llegar a una cota superior de la forma:

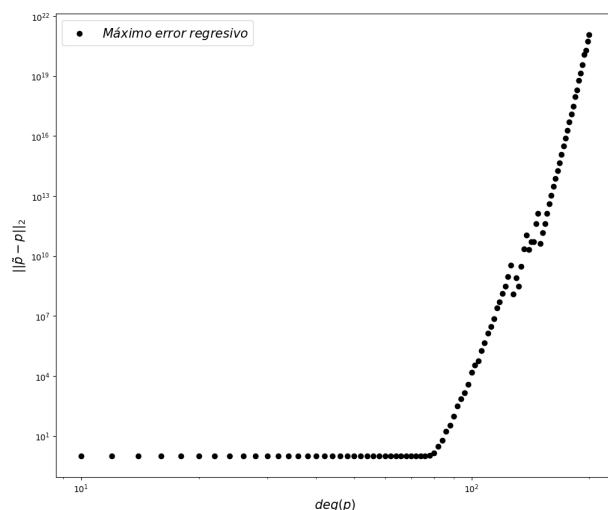
$$\|\tilde{p} - p\|_2 \leq 8n^2(n^2 + n)^{1/2}O(u)\|p\|_2^2 \quad (14)$$

Ahora, dados estos resultados analíticos procedemos a realizar simulaciones numéricas. En estas, la metodología utilizada fue la siguiente: Partiendo de un polinomio  $p(x) = T_n(x) + \sum_{k=0}^{n-1} a_k T_k(x)$  calculamos sus raíces  $\{\lambda_1, \dots, \lambda_n\}$  a partir de su matriz colega  $C_T$ . Posteriormente calculamos el error regresivo al tomar el polinomio  $\tilde{p}(x) = \prod_{k=1}^n (x - \lambda_k)$  y expandirlo en la base de Chebyshev con ayuda de las matrices de cambio de base<sup>7</sup>. Esto fue realizado en dos experimentos numéricos en los que se estudio la dependencia por separado de este error regresivo con respecto a  $\|p\|_2$  y el grado del polinomio  $p(x)$ .



**Figure 6:** Máximos errores regresivos obtenidos para cada una de las 9 muestras de 1000 polinomios aleatorios de grado 10 con norma 2 fija igual a  $10^k$ . Se utilizan en ambos ejes una escala logarítmica.

En la figura 6 se graficó el error regresivo máximo obtenido para cada una de las 9 muestras de 1000 polinomios aleatorios de grado 10 en función de la norma de los polinomios. Lo primero que se evidencia es la consistencia con la tendencia predicha por la ecuación 13. Por otro lado, en la figura 7 se graficó el error regresivo máximo obtenido para cada una de las 85 muestras de 1000 polinomios aleatorios de norma-2 igual a 1 en función del grado de los polinomios. Si bien para  $\deg(p) > 100$  se evidencia una tendencia acorde a 14, para polinomios de grado menor el error regresivo tiene un comportamiento aleatorio. Esto último se debe a errores inherentes del sistema de punto flotante de los métodos numéricos implementados.



**Figure 7:** Máximos errores regresivos obtenidos para cada una de los 95 muestras de 1000 polinomios aleatorios de grado  $n$  con norma-2 fija igual a 1. Se utilizan en ambos ejes una escala logarítmica.

**Conclusión:** Retornando a la pregunta inicial, vemos que un simple cambio de base al espacio de polinomios de Chebyshev, no solo permite calcular las raíces a partir de un problema de álgebra lineal (cálculo de valores propios), sino que soluciona el problema de mal condicionamiento de la búsqueda de las raíces. Más aún, una observación que podemos realizar, es que a partir de un problema tan básico como lo es calcular raíces se pueden encontrar fenómenos interesantes que vale la pena estudiar en futuros trabajos.

## Referencias

- [1] Arbenz, P (2018). Numerical Methods for Solving Large Scale Eigenvalue Problems Lecture Notes.
- [2] Boyd, J. P. (2002). Computing zeros on a real interval through Chebyshev expansion and polynomial rootfinding. SIAM Journal on Numerical Analysis, 40(5), 1666-1682.
- [3] Burden, R. L., & Faires, J. D. (2010). Numerical analysis. Cengage learning.
- [4] Good, I. J. (1961). The colleague matrix, a Chebyshev analogue of the companion matrix. The Quarterly Journal of Mathematics, 12(1), 61-68.
- [5] Noferini, V., & Pérez, J. (2017). Chebyshev rootfinding via computing eigenvalues of colleague matrices: when is it stable?. Mathematics of Computation, 86(306), 1741-1767.
- [6] Trefethen, L. N. (2011). Six myths of polynomial interpolation and quadrature. Mathematics Today, 47(4), 184-188.
- [7] Trefethen, L. N. (2019). Approximation Theory and Approximation Practice, Extended Edition. Society for Industrial and Applied Mathematics.
- [8] Walsh, J. L. The analogue for maximally convergent polynomials of Jentzsch's theorem, Duke Math. J. 26 (1959), 605-616.